

CLAIMS

1. A language model generation and accumulation apparatus that generates and accumulates language models for speech recognition, the apparatus comprising:

5 a higher-level N-gram language model generation and accumulation unit operable to generate and accumulate a higher-level N-gram language model that is obtained by modeling each of a plurality of texts as a sequence of words that includes a word string class having a specific linguistic property; and

10 a lower-level N-gram language model generation and accumulation unit operable to generate and accumulate a lower-level N-gram language model that is obtained by modeling a sequence of words within the word string class.

15 2. The language model generation and accumulation apparatus according to Claim 1,

wherein the higher-level N-gram language model generation and accumulation unit and the lower-level N-gram language model generation and accumulation unit generate the respective language
20 models, using different corpuses.

3. The language model generation and accumulation apparatus according to Claim 2,

25 wherein the lower-level N-gram language model generation and accumulation unit includes

a corpus update unit operable to update the corpus for the lower-level N-gram language model, and

30 the lower-level N-gram language model generation and accumulation unit updates the lower-level N-gram language model based on the updated corpus, and generates the updated lower-level N-gram language model.

4. The language model generation and accumulation apparatus according to Claim 1,

wherein the lower-level N-gram language model generation and accumulation unit analyzes the sequence of words within the word string class into one or more morphemes that are smallest language units having meanings, and generates the lower-level N-gram language model by modeling each sequence of said one or more morphemes in dependency on said word string class.

5. The language model generation and accumulation apparatus according to Claim 1,

wherein the higher-level N-gram language model generation and accumulation unit substitutes the word string class with a virtual word, and then generates the higher-level N-gram language model by modeling a sequence made up of said virtual word and the other words, said word string class being included in each of the plurality of texts analyzed into morphemes.

6. The language model generation and accumulation apparatus according to Claim 1,

wherein the lower-level N-gram language model generation and accumulation unit includes

an exception word judgment unit operable to judge whether or not a specific word out of the words that appear in the word string class should be treated as an exception word, based on a linguistic property of said specific word, and divides the exception word into (i) a syllable that is a basic phonetic unit constituting a pronunciation of said word and (ii) a unit that is obtained by combining syllables based on a result of said judgment, said exception word being a word not being included as a constituent word of the word string class, and

the language model generation and accumulation apparatus

further comprises

a class dependent syllable N-gram generation and accumulation unit operable to generate class dependent syllable N-grams by modeling a sequence made up of the syllable and the unit obtained by combining syllables and by providing a language likelihood to said sequence in dependency on either the word string class or the linguistic property of the exception word, and accumulate said generated class dependent syllable N-grams, said language likelihood being a logarithm value of a probability.

7. The language model generation and accumulation apparatus according to Claim 1, further comprising

a syntactic tree generation unit operable to perform morphemic analysis as well as syntactic analysis of a text, and generate a syntactic tree in which said text is structured by a plurality of layers, focusing on a node that is on said syntactic tree and that has been selected on the basis of a predetermined criterion,

wherein the higher-level N-gram language model generation and accumulation unit generates the higher-level N-gram language model for syntactic tree, using a first subtree that constitutes an upper layer from the focused node, and

the lower-level N-gram language model generation and accumulation unit generates the lower-level N-gram language model for syntactic tree, using a second subtree that constitutes a lower layer from the focused node.

8. The language model generation and accumulation apparatus according to Claim 7,

wherein the lower-level N-gram language model generation and accumulation unit includes

a language model generation exception word judgment unit

operable to judge a specific word appearing in the second subtree as an exception word based on a predetermined linguistic property, said exception word being a word not being included as a constituent word of any subtrees, and

5 the lower-level N-gram language model generation and accumulation unit generates the lower-level N-gram language model by dividing the exception word into (i) a syllable that is a basic phonetic unit constituting a pronunciation of said word and (ii) a unit that is obtained by combining syllables, and then by modeling a
10 sequence made up of the syllable and the unit obtained by combining syllables in dependency on a location of the exception word in the syntactic tree and on the linguistic property of said exception word.

15 9. The language model generation and accumulation apparatus according to Claim 1, further comprising

 a syntactic tree generation unit operable to perform morphemic analysis as well as syntactic analysis of a text, and generate a syntactic tree in which said text is structured by a
20 plurality of layers, focusing on a node that is on said syntactic tree and that has been selected on the basis of a predetermined criterion,

 wherein the higher-level N-gram language model generation and accumulation unit generates the higher-level N-gram language
25 model, using a first subtree that constitutes a highest layer of the syntactic tree, and

 the lower-level N-gram language model generation and accumulation unit categorizes each subtree constituting a layer lower than a second layer based on a positioning of said each
30 subtree when included in the upper layer, and generates the lower-level N-gram language model by use of each of the categorized subtree.

10. The language model generation and accumulation apparatus according to Claim 9,

wherein the lower-level N-gram language model generation and accumulation unit includes

a language model generation exception word judgment unit operable to judge, as an exception word, a specific word appearing in any subtrees in a layer lower than the second layer based on a predetermined linguistic property, said exception word being a word not being included as a constituent word of any subtrees, and

the lower-level N-gram language model generation and accumulation unit divides the exception word into (i) a syllable that is a basic phonetic unit constituting a pronunciation of said word and (ii) a unit that is obtained by combining syllables, and generates the lower-level N-gram language model by modeling a sequence made up of the syllable and the unit obtained by combining syllables in dependency on a position of the exception word in the syntactic tree and on the linguistic property of said exception word.

11. The language model generation and accumulation apparatus according to Claim 1,

wherein the higher-level N-gram language model generation and accumulation unit generates the higher-level N-gram language model in which each sequence of N words including the word string class is associated with a probability at which said each sequence of words occurs.

12. The language model generation and accumulation apparatus according to Claim 1,

wherein the lower-level N-gram language model generation and accumulation unit generates the lower-level N-gram language model by associating each N-long chain of words constituting the

word string class with a probability at which said each chain of words occurs.

13. A speech recognition apparatus that recognizes a speech
5 which is a sequence of uttered words, using the following:

a higher-level N-gram language model that is obtained by modeling each of a plurality of texts as a sequence of words that includes a word string class having a specific linguistic property; and

10 a lower-level N-gram language model that is obtained by modeling a sequence of words within the word string class.

14. A speech recognition apparatus that recognizes a sequence of uttered words, comprising

15 a language model generation and accumulation apparatus that generates and accumulates language models for speech recognition,

wherein the language model generation and accumulation apparatus includes:

20 a higher-level N-gram language model generation and accumulation unit operable to generate and accumulate a higher-level N-gram language model that is obtained by modeling each of a plurality of texts as a sequence of words that includes a word string class having a specific linguistic property; and

25 a lower-level N-gram language model generation and accumulation unit operable to generate and accumulate a lower-level N-gram language model that is obtained by modeling a sequence of words within the word string class, and

30 the speech recognition apparatus recognizes the speech by use of the higher-level N-gram language model that is accumulated by the higher-level N-gram language model generation and accumulation unit and the lower-level N-gram language model that is accumulated by the lower-level N-gram language model

generation and accumulation unit.

15. The speech recognition apparatus according to Claim 14,
wherein the higher-level N-gram language model generation
5 and accumulation unit and the lower-level N-gram language model
generation and accumulation unit generate the respective language
models, using different corpuses, and

the speech recognition apparatus recognizes the speech by
use of the higher-level N-gram language model and the lower-level
10 N-gram language model that have been respectively built using the
different corpuses.

16. The speech recognition apparatus according to Claim 15,
wherein the lower-level N-gram language model generation
15 and accumulation unit includes

a corpus update unit operable to update the corpus for the
lower-level N-gram language model,

the lower-level N-gram language model generation and
accumulation unit updates the lower-level N-gram language model
20 based on the updated corpus, and generates the updated lower-level
N-gram language model, and

the speech recognition apparatus recognizes the speech by
use of the updated lower-level N-gram language model.

25 17. The speech recognition apparatus according to Claim 14,
wherein the lower-level N-gram language model generation
and accumulation unit analyzes the sequence of words within the
word string class into one or more morphemes that are smallest
language units having meanings, and generates the lower-level
30 N-gram language model by modeling each sequence of said one or
more morphemes in dependency on said word string class, and

the speech recognition apparatus recognizes the speech by

use of the lower-level N-gram language model that has been modeled as the sequence of said one or more morphemes.

18. The speech recognition apparatus according to Claim 14,

5 wherein the higher-level N-gram language model generation and accumulation unit substitutes the word string class with a virtual word, and then generates the higher-level N-gram language model by modeling a sequence made up of said virtual word and the other words, said word string class being included in each of the plurality of texts analyzed into morphemes, and

10 the speech recognition apparatus recognizes the speech by use of the higher-level N-gram language model that has been modeled as the sequence made up of the virtual word and the other words.

15 19. The speech recognition apparatus according to Claim 18,

wherein the lower-level N-gram language model generation and accumulation unit includes

20 an exception word judgment unit operable to judge whether or not a specific word out of the words that appear in the word string class should be treated as an exception word, based on a linguistic property of said specific word, and divides the exception word into (i) a syllable that is a basic phonetic unit constituting a pronunciation of said word and (ii) a unit that is obtained by combining syllables based on a result of said judgment, said exception word being a word not being included as a constituent word of the word string class,

25 the language model generation and accumulation apparatus further comprises

30 a class dependent syllable N-gram generation and accumulation unit operable to generate class dependent syllable N-grams by modeling a sequence made up of the syllable and the

unit obtained by combining syllables and by providing a language likelihood to said sequence in dependency on either the word string class or the linguistic property of the exception word, and accumulate said generated class dependent syllable N-grams, said language likelihood being a logarithm value of a probability, and
5 the speech recognition apparatus recognizes the speech by use of the class dependent syllable N-grams.

20. The speech recognition apparatus according to Claim 19,
10 wherein the language model generation and accumulation apparatus further comprises

a syntactic tree generation unit operable to perform morphemic analysis as well as syntactic analysis of a text, and generate a syntactic tree in which said text is structured by a plurality of layers, focusing on a node that is on said syntactic tree
15 and that has been selected on the basis of a predetermined criterion,

wherein the higher-level N-gram language model generation and accumulation unit generates the higher-level N-gram language model for syntactic tree, using a first subtree that constitutes an
20 upper layer from the focused node, and

the lower-level N-gram language model generation and accumulation unit generates the lower-level N-gram language model for syntactic tree, using a second subtree that constitutes a lower
25 layer from the focused node, and

the speech recognition apparatus comprises:

an acoustic processing unit operable to generate feature parameters from the speech;

a word comparison unit operable to compare a pronunciation
30 of each word with each of the feature parameters, and generate a set of word hypotheses including an utterance segment of said each word and an acoustic likelihood of said each word; and

a word string hypothesis generation unit operable to generate a word string hypothesis from the set of word hypotheses with reference to the higher-level N-gram language model for syntactic tree and the lower-level N-gram language model for syntactic tree,
5 and generate a result of the speech recognition.

21. The speech recognition apparatus according to Claim 20,
wherein the lower-level N-gram language model generation and accumulation unit includes

10 a language model generation exception word judgment unit operable to judge a specific word appearing in the second subtree as an exception word based on a predetermined linguistic property, said exception word being a word not being included as a constituent word of any subtrees,

15 the lower-level N-gram language model generation and accumulation unit generates the lower-level N-gram language model by dividing the exception word into (i) a syllable that is a basic phonetic unit constituting a pronunciation of said word and (ii) a unit that is obtained by combining syllables, and then by modeling a
20 sequence made up of the syllable and the unit obtained by combining syllables in dependency on a location of the exception word in the syntactic tree and on the linguistic property of said exception word, and

the word string hypothesis generation unit generates the
25 result of the speech recognition.

22. The speech recognition apparatus according to Claim 14,
wherein the language model generation and accumulation apparatus further comprises

30 a syntactic tree generation unit operable to perform morphemic analysis as well as syntactic analysis of a text, and generate a syntactic tree in which said text is structured by a

plurality of layers, focusing on a node that is on said syntactic tree and that has been selected on the basis of a predetermined criterion,

wherein the higher-level N-gram language model generation and accumulation unit generates the higher-level N-gram language model, using a first subtree that constitutes a highest layer of the syntactic tree,

the lower-level N-gram language model generation and accumulation unit categorizes each subtree constituting a layer lower than a second layer based on a positioning of said each subtree when included in the upper layer and generates the lower-level N-gram language model by use of each of the categorized subtree, and

the speech recognition apparatus recognizes the speech by use of the higher-level N-gram language model that has been generated using the first subtree and the lower-level N-gram language model that has been generated using said each subtree constituting a layer lower than the second layer.

23. The speech recognition apparatus according to Claim 22, wherein the lower-level N-gram language model generation and accumulation unit includes

a language model generation exception word judgment unit operable to judge, as an exception word, a specific word appearing in any subtrees in a layer lower than the second layer based on a predetermined linguistic property, said exception word being a word not being included as a constituent word of any subtrees,

the lower-level N-gram language model generation and accumulation unit divides the exception word into (i) a syllable that is a basic phonetic unit constituting a pronunciation of said word and (ii) a unit that is obtained by combining syllables, and generates the lower-level N-gram language model by modeling a sequence

made up of the syllable and the unit obtained by combining syllables in dependency on a position of the exception word in the syntactic tree and on the linguistic property of said exception word, and

the speech recognition apparatus recognizes the speech by use of the higher-level N-gram language model that does not include the exception word and the lower-level N-gram language model that includes the exception word.

24. The speech recognition apparatus according to Claim 14, wherein the higher-level N-gram language model generation and accumulation unit generates the higher-level N-gram language model in which each sequence of N words including the word string class is associated with a probability at which said each sequence of words occurs, and

the speech recognition apparatus comprises a word string hypothesis generation unit operable to evaluate a word string hypothesis by multiplying each probability at which said each sequence of N words including the word string class occurs.

25. The speech recognition apparatus according to Claim 14, wherein the lower-level N-gram language model generation and accumulation unit generates the lower-level N-gram language model by associating each N-long chain of words constituting the word string class with a probability at which said each chain of words occurs, and

the speech recognition apparatus comprises a word string hypothesis generation unit operable to evaluate a word string hypothesis by multiplying each probability at which said each sequence of N words inside the word string class occurs.

26. A language model generation method for generating language

models for speech recognition, comprising:

a higher-level N-gram language model generation and accumulation step of generating and accumulating a higher-level N-gram language model that is obtained by modeling each of a plurality of texts as a sequence of words that includes a word string class having a specific linguistic property; and

a lower-level N-gram language model generation and accumulation step of generating and accumulating a lower-level N-gram language model that is obtained by modeling a sequence of words within the word string class.

27. A speech recognition method for recognizing a speech which is a sequence of uttered words, using the following:

a higher-level N-gram language model that is obtained by modeling each of a plurality of texts as a sequence of words that includes a word string class having a specific linguistic property; and

a lower-level N-gram language model that is obtained by modeling a sequence of words within the word string class.

28. The speech recognition method comprising:

a step of categorizing each word string having a specific linguistic property as a word string class, and providing, to said each word string, a language likelihood which is a logarithm value of a probability, by use of class dependent word N-grams that are obtained by modeling said word string class in dependency on said word string class based on a linguistic relationship between words constituting said word string class;

a step of analyzing a text into a word and the word string class, and providing, to a sequence of said word and the word string class, a language likelihood which is a logarithm value of a probability, by use of class N-grams that are obtained by modeling said sequence of the word and the word string class based on a linguistic relationship;

and

a step of (i) comparing features parameters extracted from a series of speeches with a pronunciation as well as an acoustic characteristic of each word and generating a set of word hypotheses including an utterance segment of said each word and an acoustic likelihood of said each word, (ii) generating a word string hypothesis from said set of word string hypotheses with reference to the class N-grams and the class dependent word N-grams, and (iii) outputting a result of the speech recognition.

10

29. A program for performing a language model generation method that is intended for generating a language model for speech recognition, the program causing a computer to execute the following steps:

15

a higher-level N-gram language model generation and accumulation step of generating and accumulating a higher-level N-gram language model that is obtained by modeling each of a plurality of texts as a sequence of words that includes a word string class having a specific linguistic property; and

20

a lower-level N-gram language model generation and accumulation step of generating and accumulating a lower-level N-gram language model that is obtained by modeling a sequence of words within the word string class.

25

30. A program for performing a speech recognition method that is intended for recognizing a sequence of uttered words, the program causing a computer to execute a speech recognition step that is performed by use of the following:

30

a higher-level N-gram language model that is obtained by modeling each of a plurality of texts as a sequence of words that includes a word string class having a specific linguistic property; and
a lower-level N-gram language model that is obtained by

modeling a sequence of words within the word string class.